

論文

統制的発表語彙サイズテストの分析

¹ 三 根 浩 ² 今 井 由美子¹ 同志社女子大学・学芸学部・情報メディア学科・教授² 同志社女子大学・表象文化学部・英語英文学科・准教授

Analysis of a Vocabulary Size Test on Controlled Productive Ability

¹ Hiroshi Mine ² Yumiko Imai¹ Department of Information and Media, Faculty of Liberal Arts,
Doshisha Women's College of Liberal Arts,
Professor² Department of English, Faculty of Culture and Representation,
Doshisha Women's College of Liberal Arts,
Associate Professor

Abstract

Vocabulary knowledge is an important part of measuring EFL students' overall language ability. Generally vocabulary levels are assessed by measuring receptive knowledge. However, to train language learners to become genuinely capable communicators, it is essential to measure not only receptive knowledge but also productive knowledge. To accomplish this, Laufer and Nation (1995) developed the Vocabulary Size Test of Controlled Productive Ability (VLT-CPA). Unfortunately, this test has proved difficult to use with Japanese learners. This paper investigates the possibility of measuring productive vocabulary size of Japanese EFL learners using the VLT-CPA. Results show that, 1) participants' VLT-CPA scores correspond to the 45% mark commonly found when using the Vocabulary Levels Test (VLT), thereby reconfirming the results of Waring (1997). Further, 2) productive vocabulary size of the CELT High-score group and Low-score group examined by the VLT-CPA show an almost 30% difference. Thus we determine that the VLT-CPA can be used with Japanese EFL learners, and be regarded as an accurate indicator of productive vocabulary knowledge. Finally, 3) the High-score group obtained scores of about 60% in the VLT-CPA at the 2,000 word level, confirming the importance of encouraging students to strive for a productive vocabulary of at least 2,000 words, in order to become successful language users.

キーワード：EFL、語彙サイズ、発表の語彙、受容の語彙

1. はじめに

英語教育において語彙力の測定は重要な指標の一つであり、指導者は学習者のレベルを把握し、目的に合わせた指導法を取り入れ、学習者が学習を通しさらに語彙力を伸ばしていくための支援を行うべきである。語彙サイズや語彙レベルの測定では一般的に受容的語彙がテストされるが、実用的な語学力を育成するためには発表的語彙力も重要である。

語彙学習や語彙サイズ測定のために活用されている学習用語彙リストとしては、A teacher's word book of 30,000 words (Thorndike and Lorge, 1944)、General Service List (West, 1953)、Academic Word List (Coxhead, 2000)、JACET8000 (大学英語教育学会基本語改訂委員会, 2003) などがあり、近年では、時代の流れに伴い A New General Service List (1.01) (Browne, Culligan, and Phillips, 2013b) および A New Academic Word List (Browne, Culligan, and Phillips, 2013a) が編纂され発表された。語彙力を測定するテストは、個々の語彙リストからグレード別の単語サンプルを採り、その中の既知語の割合を推定することにより全体的な語彙サイズを算出する。

語彙テストには、受容的語彙テスト (単語の文字形式、つまり単語の綴りを提示し意味を選択させ知識をチェックさせる再認形式) と発表的語彙テスト (単語や文章を書かせる再生形式) とがある。受容的語彙力を測定するものとして、望月 (1998) による語彙サイズテスト、Schmitt や Nation による語彙レベルテスト (Vocabulary Levels Test; VLT, Nation, 2001; Schmitt, 2000; Schmitt, Schmitt, and Clapham, 2001)、Meara and Buxton (1987) のチェックリスト・テストなどが、また発表的語彙力を測定するテストとして Laufer and Nation (1995) の語彙頻度プロフィール法や統制的発表語彙サイズテスト (Vocabulary Size Test of Controlled Productive Ability: VLT-CPA, Laufer and Nation, 1999) などが代表的である。一般的に、日本人 EFL 学習者が発表的語彙力のテストを受験するのはむずかしすぎると考えられてきた。一方で、実用的な語学力を育成する上で発表的側面からの語彙力を知ることが必要であり、EFL 分野を中心に発表語彙テストに関する実践的研究が行われるようになってきている。

Fitzpatrick and Clenton (2010) は、日本人 EFL 学習者を対象に Lex30 という発表語彙テストの妥当性の検証に挑戦した。Lex30 は、単語連想課題により解答者から語彙を誘発するテストであり、英語の発表語出現頻度に基づい

て採点される点が単なる連想テストとの違いである。課題は30個の手がかり単語から構成される。課題語は、1) 出現頻度が最初の1,000語であり、2) 強い初頭反応を出さないこと、3) 誘発された反応語が最初の1,000語に含まれないこと、といった3条件を満たす必要があり Edinburgh Association Thesaurus の反応データ (<http://www.eat.rl.ac.uk/>) が基準として用いられた。

2. 先行研究

2.1 VLT

語彙レベルテスト (VLT) は、Nation (1990)、Schmitt (2000) および Schmitt, et al. (2001) により1980年代から開発が進められ、現在は Schmitt が著作権を有している。テストは、2,000語 (ワード・ファミリー) レベル、3,000語レベル、5,000語レベル、10,000語レベル、Academic 語レベルの5種類から構成されている。Nation (1990) では3つの定義文と6単語の選択肢から構成された設問が6組ずつあり18点満点であったが、Schmitt, et al. (2001) では10組ずつ用意され30点満点となった。これらのレベルの単語は、Thorndike and Lorge (1944) の語彙リストから、West (1953) の General Service List と Kucera and Francis (1967) の頻度表を参照して層別抽出されている。このうち、2,000語レベルの単語は最初の1,000語と次の1,000語から1:2の比率で抽出され、Academic 語レベルの単語は Coxhead (2000) の Academic 語からサンプリングされている。

三根・枝澤・吉村・今井・布施・平岩 (2006) は、英語標準テスト (CELT) リスニング部門の得点と、語彙レベルテストと語彙サイズテストおよび語彙学習方略の関連性を検討した。その結果、語彙力とリスニング力が一貫して有意な相関係数を示し、リスニング力の育成に語彙力の増強が効果的であることが示された。さらに、学習方略の因子分析の結果、補償方略がアカデミック語の学習と関連性の高いこと、単語を直接的に記憶しようとする方略は高い語彙力やリスニング力を持つ学習者にとって適切でないこと、音声を日本語に置き換える理解方略が基本語彙の増強に結びつかないことなどを明らかにした (p.67)。

枝澤・今井・古荘・布施・三根 (2007) は日本人 EFL 学習者に2,000語か Academic 語の辞書調べを課すことが CELT および TOEIC のリスニング得点に及ぼす効果を検討した。その結果、辞書調べの課題によるグループ間にリスニング力得点の伸びに有意な差が見られなかった。これ

に対し、TOEIC 得点が大幅に上昇した学習者は、3,000語レベルの得点の増加がもっとも大きく、TOEIC 得点が減少した学習者では、学習初期において2,000語レベルの得点が24点に到達していなかったことが明らかにされた(p. 61)。この結果は、EFL 学習者にできるだけ早く2,000語レベルをマスターさせ、続いて3,000語レベルを学習することの重要性を示唆した。

2.2 VLDT

ディクテーション形式による語彙レベルディクテーションテスト (Vocabulary Levels Dictation Test; VLDT, Fountain and Nation, 2000) はヒアリングという受容的スキルと聞き取った単語や文章を書き取るという発表的スキルにより構成されている。このテストの解答者は、一連の音の流れから単語を一つずつ文節化し、順次に記憶し、解答用紙の上にセンテンスを再構成することが求められる。VLDT は Test-A~D の4つの並行テストがあり、各テストは5つのパラグラフからなる文章で構成されている。ターゲット語は、導入パラグラフ (基本語彙1~500語)、第1パラグラフ (基本語彙501~1,000語)、第2パラグラフ (1,001~2,000語)、第3パラグラフ (2,001~3,000語)、第4パラグラフ (4,000~6,000語) にグレードに分けて配置されている。採点の対象となるターゲット語は各パラグラフに20語ずつ含まれ、導入パラグラフを除き、第1パラグラフから第4パラグラフまで各20点ずつ、合計80点満点で採点される。Fountain and Nation は、これらの4並行テストの相互相関は0.95以上、語彙レベルテストとの相関は0.78、Grammar test との相関は0.70以上であり、高い信頼性と妥当性を有すると報告している。

今井・三根 (2010) は日本人 EFL 学習者の発表的語彙知識を測るために、VLDT の適用可能性とテストの信頼性を検証した。VLDT の Test-A と Test-B を原版のまま実施し両テストの相関係数と平均得点の差異を検討した結果、Test-A と Test-B の相関係数は0.79あったものの、フレーズの中に含まれる単語数が多いことによる Test-A の課題の困難さが明らかになった。そこで、VLDT の困難性を解決するために、Test-A のテキスト中、6語以上の語数からなるフレーズを発音の切れ目で分割することにより修正を加え、単語レベルでの誤聴分析を試みた。修正版では、長いフレーズを分割した結果、ポーズの前後の単語における正答率が向上した。VLDT は十分に高い信頼性を示したが、調査協力者の平均正答率が非常に低いことは語彙力よりも音韻ループの容量に制約を受けていた可能性

を示唆した。VLDT は、フレーズに含まれる単語数が多いことにより日本人 EFL 学習者にとって難しい課題になった可能性があり、VLDT では語彙力よりも記憶スパンを測定している比重が大きいと考えられることを指摘した。

今井・三根 (2011) は CELT を外的基準として VLDT の語彙力テストとしての信頼性と妥当性の検証を試みた。また、VLT と VLDT の成績を語彙サイズに換算することで受容語彙力と発表語彙力の比較を試みた。VLT から推定される語彙サイズに比較して、VLDT の Test-A、B、D では48.4%~50.0%であったが、VLDT の Test-C の得点は、VLT に対し41.1%にしか達せず期待を裏切る結果を示した。結論として、VLDT 総得点が十分な信頼性を示した一方で、VLDT-C は基準関連妥当性を欠いていること、CELT による英語力の高い学習者が受容的語彙サイズから期待されるよりも高い発表的語彙サイズを持つ可能性が示されたこと、CELT による英語力の低い学習者の VLDT 得点が全般的に低かったことなどが示された(p. 92)。

今井・三根 (2012) は、日本人 EFL 学習者を対象に、VLDT の語彙テストとしての信頼性について再検証を試みた。VLDT の4つの並行テストのうち Test-C の難易度を、今井・三根 (2010) の Test-A (再修正版) と同様に音声的・文法的に切り分け可能な箇所ではポーズを追加し修正することにより、並行テスト間のテスト得点の同質性を検証した。その結果、日本人 EFL 学習者にとって発表的語彙力テストの側面を有する VLDT は、語彙力の指標として VLDT のパラグラフ別の得点ではなく全項目の総得点を採用することが適切であることを再確認した。

その結果、EFL 学習者において、ディクテーション形式の語彙テストでは、1) 高レベルの語彙のヒアリング、2) 難しいスペリング、3) 複数形の処理、4) 長い音節数の語彙のヒアリングが困難であることが明らかにされた。とくに、複数形の困難性は、EFL 学習者がディクテーションにおいて文脈的・文法的処理を行わず、単語を逐語的に処理していることを示唆した (p. 53)。

2.3 VLT-CPA

統制的発表語彙サイズテスト (VLT-CPA, Laufer and Nation, 1995) は、受容語彙を測る VLT (Nation, 1990) をベースに開発された、発表的語彙の測定を目的としたテストである。一般的に、このテストは Productive Levels Test と略称されているが、ここでは VLT を元にした統制

的発表能力テストであることを明確にするため VLT-CPA と略称した。VLT-CPA は C-test に似たテスト形式で、1つの語彙は1つのセンテンスの中で数文字の頭字手がかりに基づいてテストされる。このテストの解答者は、センテンスの文脈から正しい語彙を推測し頭字手がかりに従ってスペリングを完成することが求められる。2,000語、3,000語、University 語、5,000語、10,000語の5レベルがある。University 語は Xue and Nation (1984) の University Word List からサンプリングされている。各レベルには Nation (1990) の VLT で用いられた語彙をターゲット語とする18個の問題文が作られた。18点満点中少なくとも15点以上あることが望ましく (Nation, 1999, p. 140)、各レベルで80%以上の成績を得ることは、発表語彙として十分に使用できることを示している。VLT-CPA は Version C と呼ばれるオリジナルテストに加え、3つの別バージョンのテスト形式の中から、Laufer and Nation (1999) による信頼性分析の結果に基づいて Version 2が作成され、2つの並行テストが採用された。

2.4 日本人 EFL 学習者における VLT-CPA

Waring (1997) は、受容的・発表的語彙知識の性質に着目した。日本人 EFL 学習者に対し受容的語彙知識を測定する VLT と統制的発表語彙サイズテスト (VLT-CPA) の2,000語、3,000語および5,000語レベルに独自に作成した1,000語レベルを加えたテストを実施し、学習者の語彙知識がどれ位あるかということだけでなく、受容的語彙知識が発表的語彙知識を何%ほど上回るのかを調査した。その結果は受容語彙と発表語彙の比が44.8% (Table 1, p. 60) で、「受容的語彙知識は発表的語彙知識よりも大きいことを明確に示すデータを見いだした」と報告した (p. 63)。語彙を受容的に理解する場合は、語の文字形式あるいは音声形式と意味とを合致させるスキルが必要となる。一方で語彙を発表的に使用する場合は、語の意味、スペリングや発音、文法、社会的言語使用における知識など、受容的使用と比較すると格段に多くのスキルが求められる。

2.5 発表語彙テストの問題点

Fitzpatrick and Clenton (2010) は、VLT タイプの発表語彙テストの問題点として、1) 学習者のレキシコンが成長するほど、語彙サイズを外挿法で推定するための単語を十分に用意できなくなること、2) 文脈を通して単語を誘発する際に受容的知識も使用されるため発表的知識のみに特化した結論を導けないことを挙げた。後者については、とくにディクテーションを介在させる VLDT の結果に関して多くの研究者から批判を受ける理由の1つになっている。一方、Lex30は短時間で実施できるが、他の発表語彙テストとの相関は有意であるが数値が低いこと、ネイティブ話者とそうでない学習者を判別できないことが問題点として指摘された。

2.6 目的

本研究の目的は、日本人 EFL 学習者を対象として、Laufer and Nation (1999) が開発した統制的発表語彙サイズテスト (VLT-CPA) の適用可能性を検証することである。

3. 方法

3.1 研究協力者

研究協力者は英語専攻の大学2年生で秋学期に週1回90分のリスニング科目を受講した女子学生である。秋学期末(1月)に、CELT (Listening Part Form A ; 30分)、VLT (Version 2 ; 30分)、VLDT (Test D ; 15分)、VLT-CPA (Version 1 ; 30分) の4つのテストを受験した113名を調査の対象とした。研究協力者には、書面にて調査への協力を確認した。CELT の平均点は32.37点であったため、32点以下を CELT-Low 群 (56名) とし33点以上を CELT-High 群 (57名) とした。

3.2 手続き

VLT-CPA は Laufer and Nation (1999, pp. 46-48) にある Version 1を使用した。Laufer and Nation には解答欄が

Table 1 CELT、VLT-size、VLDT-size と CPA-size の相関係数行列

	VLT	VLDT	VLT-CPA
CELT	0.56	0.62	0.54
VLT-Size	1.00	0.50	0.65
VLDT-Size	0.50	1.00	0.51
CPA-Size	0.65	0.51	1.00

示されていなかったが、採点の便宜をはかるため解答欄を右側に設け、下線の長さは解答すべき語の長さに関係なく統一し、文中で頭字手がかりとして与えられている文字を置いた（Appendix 参照）。10,000語レベルを除く4つのレベルを使用し、解答時間は30分とした。解答は原文に記載されていなかったため、筆者らが解答し英語母語話者に確認を依頼した。センテンスの文脈から正しい語彙を推測しスペルアウトする発表的語彙力を測定するため、複数形や動詞の過去形、過去分詞形などスペリングが正確に綴られている場合のみ正解とした。

4. 結 果

4.1 語彙サイズによる比較

Table 1は各種の測定値を比較するためにVLT、VLDT、VLT-CPAを望月（2003、p. 211）に準じて語彙サイズに換算し、CELTの得点とともに相関係数を求めたものである。その結果、CELT、VLT、VLDT、VLT-CPAはプラス0.5以上ですべて有意性水準1%未満の相関係数を示した。

Table 2はCELTの平均得点に基づいて研究協力者を高得点（High）群57名と低得点（Low）群56名に分け、VLT、VLDT、およびVLT-CPAにおける語彙サイズの差を t 検定（ $df = 111$ 、以下同様）で比較した。すべてのテストの下位尺度において高得点群の方が統計的に1%水準で有意に高い語彙得点を示した。

VLT、VLDT、VLT-CPAの3種の語彙テストの得点を語彙サイズに変換した結果、受容的語彙力を測るVLTに対して、VLDTの平均語彙サイズが51%、VLT-CPAの平均語彙サイズが45%に相当する結果となった。語彙の発表

的能力を測るテストとして、ディクテーション形式のVLDTよりも記述式の統制的発表語彙サイズテストのVLT-CPAの方がより困難であることが明らかになった。とくに、CELT低得点者ではVLT-CPAはVLTサイズの42%にしか達せず、高得点者の語彙サイズよりも30%近く低い有意な差（ $t = 5.62$ ）を示した。

4.2 CELT と VLT-CPA

Table 3はCELT高得点群と低得点群におけるVLT-CPA得点の差をまとめた。VLT-CPAの2,000語、3,000語、University語、5,000語の各レベルにおいて、CELT高得点群は有意に高い得点を示した。2,000語では18点満点中、高得点群が10.9点の60%に対し、低得点群では8.3点で46%、全体として50%を超える9.6点の正答が得られた。しかしながら、CELT高得点群でも、2,000語レベルが60%にしか達していない状態では、発表語彙力が実用的レベルに達しているとはみなせない。3,000語以上の語彙レベルの平均点では、3,000語が3.8点で約21%、University語が3.6点で約20%、5,000語レベルが2.4点で約14%の正答率であった。とくに5,000語における低得点群は1.86点で10.3%の正答率にしか達しなかった。

4.3 CELT と VLT

Table 4はCELT高得点群と低得点群におけるVLT得点の差をまとめた。CELT高得点群は2,000語と3,000語レベルを合わせると30点満点中80%の正答率を超えた。Laufer（1992）は、“reader”が“non-reader”より多くなる語彙レベルは3,000語である（p. 130）と指摘したが、そのレベルをクリアしていることを示した。また、5,000語レベルにおいても低得点群の平均点は50%に近い正答率で

Table 2 CELT High 群と Low 群での語彙サイズの差

	CELT-Low	CELT-High	t	p
VLT-Size	3075	3781	-6.08	0.00
VLDT-Size	1549	1965	-5.68	0.00
CPA-Size	1299	1791	-5.62	0.00

Table 3 CELT High 群と Low 群での VLT-CPA 得点の差

	CELT-Low	CELT-High	t	p
CPA2000	8.27	10.86	-5.03	0.00
CPA3000	3.13	4.49	-3.86	0.00
CPA_UWL	2.63	4.46	-4.67	0.00
CPA5000	1.86	3.02	-4.21	0.00

Table 4 CELT High 群と Low 群での VLT 得点の差

	CELT-Low	CELT-High	<i>t</i>	<i>p</i>
VLT_2000	22.82	26.61	-5.10	0.00
VLT_3000	17.54	21.83	-5.02	0.00
VLT_Aademic	17.57	22.30	-5.12	0.00
VLT_5000	14.54	19.19	-5.27	0.00

Table 5 CELT High 群と Low 群での VLDT 得点の差

	CELT-Low	CELT-High	<i>t</i>	<i>p</i>
PARA_D1	11.54	13.63	-5.61	0.00
PARA_D2	4.98	6.40	-4.06	0.00
PARA_D3	5.04	7.30	-4.88	0.00
PARA_D4	4.71	5.98	-3.46	0.00

あった。視覚的情報を頼りに正答を導き出すことができるため、受容的語彙テストである VLT では、高得点群および低得点群の正答率は全体的に高かったといえる。

4.4 CELT と VLDT

Table 5は CELT 高得点群と低得点群における VLDT の各パラグラフの得点の差をまとめた。採点の対象となるターゲット語は第1パラグラフから第4パラグラフまで各20語ずつあり、合計80点満点で採点された。第1パラグラフ（基本語彙501～1,000語）では全体の約65%の正答率だが、第2パラグラフ（1,001～2,000語）、第3パラグラフ（2,001～3,000語）および第4パラグラフ（4,000～6,000語）における平均正答率は約25%であった。VLDTにおいて解答者は、聞こえてくる音声記憶し、解答用紙の上にスペルアウトすることが求められる。長いフレーズでは中央にある単語が記憶に残りにくいことや語彙レベルが高くなるにつれて単語のシラブル数が増えること、正しくスペリングすることや意味的・文法的処理をすることの困難性も加わり、1,000語レベルを超えると平均得点は極度に低いものとなった。その一方で、4,000～6,000語レベルの語彙力を測定する第4パラグラフにおいても第2および第3パラグラフと大差のない平均正答率に達した。第4パラグラフにターゲット語として配置された語の中には、Longman の辞書では S1または S2、W1または W2といった1,000語～2,000語にランクされている語も含まれていた。4,000～6,000語レベルのターゲット語にこのような単語が採用されたのは、70年も前の Thorndike and Lorge (1944) の語彙リストが用いられたことが一因かもしれない。当時はこの語彙表で5,000語レベルにリストされてい

た単語も、時代の推移とともに日常的になじみのある単語となってきたため、研究協力者にとっては聴覚手がかりから容易に正解に結びつけられたと推測される。

5. 考 察

本研究では Laufer and Nation (1999) が開発した統制的発表語彙サイズテスト (VLT-CPA) の EFL 学習者への適用可能性を検討した。VLT-CPA の得点は受容的語彙テストである VLT の45%に相当し、Waring (1997) の結果を再確認することとなった。また、CELT による英語力の高得点群と低得点群の間には30%もの発表語彙サイズの差が見出され、VLT-CPA が英語力の指標として妥当なものであることが確認された。高得点群が2,000語レベルにおいて約60%の得点しか得られなかったことは、EFL 学習者にとって課題が難しすぎたことを意味する。

VLT や VLDT と VLT-CPA を比較すると、2,000語レベル以外では VLT-CPA テストの正答率が非常に低いことが示された。VLT は再認テストという形により受容的スキルを測定するもので、これら3つのテストの中では最も正答率が高かった。VLDT は、音韻情報をヒアリングする受容的課題と聞き取った単語のスペリングを再生する発表的課題の2種類のスキルを測定するテストで、与えられた音韻情報を頼りにスペリングの情報を記憶に留め、意味的・文法的処理を行い、スペルアウトするというスキルが求められた。VLT-CPA は、文脈から正しい語彙を推測し、数文字の頭字手がかりに基づいてスペリングを完成させる発表的スキルを測定するテストで、3種類のテストでは困難度が高かった。この文脈においてはこの語しか

該当しないというターゲット語の頭字手がかりが与えられているものの、単語そのものが発表的語彙として身につけていなければ実用の域には至らない。発表的課題を行うためには、文脈における正しい形式を作りだす上で意味的・文法的処理も同時に求められるため、より高度な英語運用力が必要となる。投野 (2013) は「高頻度語2,000語の語彙群をきちんと発信語彙として習得することの重要性」(p. 282) を強調し、「コア語彙として高頻度語の2,000語がマスターできなければ、その英語は実用レベルで使いものにならない」(p. 283) と指摘した。2,000語レベルの語彙が未習得で英語力が低い学習者には、正しいスペリングの指導を含めた発表語彙力の育成が重要であると考えられる。本研究で採用した VLT-CPA は、2,000語レベルの発表語彙力が日本人 EFL 学習者学習者において不足していることを示した。成功的外国語学習者となる上で、2,000語レベルの発表語彙の獲得を目指すことの重要性を確認した。

注：本研究は、外国語教育メディア学会第54回（2014年度）全国研究大会（於：福岡大学）において口頭発表したものに加筆修正を加えたものである。

6. 引用文献

- Browne, C., Culligan, B., & Phillips, J. (2013a). A New Academic Word List.
<http://www.newacademicwordlist.org/>
- Browne, C., Culligan, B., & Phillips, J. (2013b). A New General Service List.
<http://www.newgeneralservicelist.org/>
- Coxhead, A. (2000). A new Academic Word List, *TESOL Quarterly*, 34, 213-238.
- 大学英語教育学会基本語改訂委員会 (2003). 『大学英語教育学会基本語リスト JACET List of 8000 Basic Words』.
- 枝澤康代・今井由美子・古荘智子・布施邦子・三根浩 (2007). 「大学生における語彙力と英語標準テストの関連性」『同志社女子大学総合文化研究所紀要』 24, 55-66.
- Fitzpatrick, T., & Clenton, J. (2010). The challenge of validation: Assessing the performance of a test of productive vocabulary. *Language Testing*, 27, 537-554.
- Fountain, R. L., & Nation, I. S. P. (2000). A vocabulary-based graded dictation test. *RELC Journal*: *Guidelines*, 1, 76-80.
- 今井由美子・三根浩 (2010). 「語彙レベルディクテーションテストの日本人学習者への適用」『JACET 関西紀要』 12, 56-67.
- 今井由美子・三根浩 (2011). 「ディクテーション形式による語彙レベルテストの妥当性の検討」*Language Education & Technology*, 48, 83-94.
- 今井由美子・三根浩 (2012). 「Vocabulary Levels Dictation Test の信頼性の再分析」『外国語教育メディア学会第52回全国研究大会論文集』 52-53.
- Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In Arnaud, P. J. L., and Bejoint, H. (Eds.), *Vocabulary and applied linguistics*. London: Macmillan. pp. 126-132.
- Laufer, B., & Nation, I. S. P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16, 307-322.
- Laufer, B., & Nation, I. S. P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16, 33-51.
- Meara, P., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests, *Language Testing*, 4, 142-151.
- 三根浩・枝澤康代・吉村満知子・今井由美子・布施邦子・平岩葉子 (2006). 「リスニングにおける語彙サイズと学習方略」『同志社女子大学総合文化研究所紀要』 23, 59-69.
- 望月正道 (1998). 「日本人学習者のための英語語彙サイズテスト」『語学教育研究所紀要』 12, 27-53.
- 望月正道 (2003). 「語彙はどうテストするのか」望月正道、相澤一美、投野由紀夫 (編) 『英語語彙の指導マニュアル』大修館書店. pp. 181-209.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Boston: Heinle and Heinle.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions

- of the vocabulary levels test. *Language Testing*, 18, 55-88.
- Thorndike, E. L., & Lorge, I. (1944). *The teachers' word book of 30,000 words*. Columbia University: Teachers College.
- 投野由紀夫 (2013). 『英語到達度指標 CEFR-J ガイドブック』 東京: 大修館書店.
- Waring, R. (1997). A comparison of the receptive and productive vocabulary sizes of some second language learners. *Immaculata; The Occasional Papers of Notre Dame Seishin University*, 1, 53-68.
- West, M. (1953). *A general service list of English words*. London: Longman, Green & Co.
- Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication*, 3, 215-229.

Appendix A Levels Test of Productive Vocabulary Version 2 (Sample Page)

A LEVELS TEST OF PRODUCTIVE VOCABULARY: Version 2 (Version C)

Complete the underlined words. The example has been done for you.

He was riding a bicycle. ※解答はすべて右側の解答欄へ記入して下さい。

The 2000-word level

1. It is the de_____ that counts, not the thought.
2. Plants receive water from the soil through their ro_____.
3. The nu_____ was helping the doctor in the operating room.
4. Since he is unskilled, he earns a low wa_____.
5. This year, long sk_____ are fashionable again.
6. Laws are based upon the principle of jus_____.
7. He is walking on the ti_____ of his toes.
8. The mechanic had to replace the mo_____ of the car.
9. There is a co_____ of the original report in the file.
10. They had to cl_____ a steep mountain to reach the cabin.
11. The doctor ex_____ the patient thoroughly.
12. The house was su_____ by a big garden.
13. The railway con_____ London with its suburbs.
14. She wan_____ aimlessly in the street.
15. The organizers li_____ the number of participants to fifty.
16. This work is not up to your usu_____ standard.
17. They sat down to eat even though they were not hun_____.
18. You must have been very br_____ to participate in such a dangerous operation.

【解答欄】

1. de_____
2. ro_____
3. nu_____
4. wa_____
5. sk_____
6. jus_____
7. ti_____
8. mo_____
9. co_____
10. cl_____
11. ex_____
12. su_____
13. con_____
14. wan_____
15. li_____
16. usu_____
17. hun_____
18. br_____